Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT

ELSEVIER

Stochastic Processes and their Applications 104 (2003) 173–192

stochastic processes and their applications

www.elsevier.com/locate/spa

Trimmed sums for non-negative, mixing stationary processes

Jon Aaronson[a,*], Hitoshi Nakada[b]

[a] *School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel*

[b] *Department of Mathematics, Keio University, Hiyoshi 3-14-1 Kohoku, Yokohama 223, Japan*

Received 15 May 2002; received in revised form 13 September 2002; accepted 18 September 2002

Abstract

We consider the effect of “trimming” ergodic sums of their maximal values on the strong law of large numbers for non-negative, non-integrable, mixing stationary processes. The results obtained are used to show the failure of the strong law of large numbers for modified continued fraction coefficients, and to study the “cusp visits” of a certain interval map.

MSC: 60F; 11K; 37A; 28D; 37E

Keywords: Trimmed sums; Infinite expectation; Asymptotics; Continued fractions; Interval maps

0. Introduction

0.1. Laws of large numbers and sum trimming

We consider non-negative, $\mathbb{R}$-valued ergodic, stationary processes $(X_1, X_2, \ldots)$. In case $E(X_1) = \infty$, there is no strong law of large numbers for the partial sums $S_n := \sum_{k=1}^{n} X_k$.

It is shown in Aaronson (1977) (see also Aaronson, 1997, Section 2.3) that if $b_n > 0$ are constants then,

$$\text{either} \quad \overline{\lim_{n\to\infty}} \frac{1}{b_n} S_n = \infty \quad \text{a.s.,} \qquad \text{or} \qquad \underline{\lim_{n\to\infty}} \frac{1}{b_n} S_n = 0 \quad \text{a.s.} \tag{0.1}$$

See Feller (1946) and Chow and Robbins (1961) for the original proofs in the i.i.d. case.

* Corresponding author.

E-mail addresses: aaro@post.tau.ac.il (J. Aaronson), nakada@math.keio.ac.jp (H. Nakada).

0304-4149/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0304-4149(02)00236-3

There may be a weak law of large numbers when $E(X_1)=\infty$. Feller (1945) showed that if $(X_1, X_2, \ldots)$ are non-negative, i.i.d. random variables, the weak law of large numbers holds in the sense that

$$\exists b(n) \text{ constants such that } \frac{S_n}{b(n)} \xrightarrow{\mathrm{P}} 1 \tag{0.2}$$

(where $\xrightarrow{\mathrm{P}}$ denotes stochastic convergence) iff $L(t) := E(X \wedge t)$ is slowly varying at ∞ (see below) and in this case $b(n) \sim nL(b(n))$.

The strong law here breaks down in a particular way: since $E(X)=\infty \Rightarrow E(b^{-1}(X)) = \infty$, we have (by the Borel–Cantelli lemma)

$$\overline{\lim_{n\to\infty}} \frac{S_n}{b(n)} \geqslant \overline{\lim_{n\to\infty}} \frac{X_n}{b(n)} = \infty \quad \text{a.s.} \tag{0.3}$$

The question arose as to whether the maximal terms of $\{X_1, \ldots, X_n\}$ are "responsible" for the failure of the strong law, particularly in view of the fact that under the additional assumption that $L(t) \sim L(t \log\log t)$ (as shown in Klass and Teicher, 1977),

$$\underline{\lim_{n\to\infty}} \frac{S_n}{b(n)} = 1 \quad \text{a.s.} \tag{0.4}$$

Mori studied strong laws for i.i.d. random variables when finitely many of these maximal terms are excluded (trimmed) from the sums S_n and characterised (in terms of the distribution of the X_k and the normalising constants) when a trimmed strong law holds (see Mori, 1976, 1977).

In this paper, we consider such trimming for dependent processes, extending a theorem of Mori's (Theorem 1.1 below) to certain continued fraction mixing processes (see below), and exhibiting Markov chains (satisfying (0.2)–(0.4)) for which it fails.

For simplicity, we restrict attention to non-negative processes, as in the general $\mathbb{R}$-valued case, there may be interaction of the positive and negative parts causing strong laws which are spurious from the viewpoint of this paper.

0.2. Regular variation

Recall (from Karamata, 1933; Bingham et al., 1987; Feller, 1966) that a measurable function $f:\mathbb{R}_+ \to \mathbb{R}_+$ is called *regularly varying* (at $\theta = 0, \infty$) if $\forall \lambda > 0$, $\exists \lim_{t\to\theta} f(\lambda t)/f(t) =: \ell(\lambda)$. In case f is regularly varying, the function ℓ is necessarily of form $\ell(\lambda) = \lambda^\alpha$ for some $\alpha \in \mathbb{R}$ which is called the *index* (of regular variation of f).

The function $f:\mathbb{R}_+ \to \mathbb{R}_+$ is called *slowly varying* at θ if it is regularly varying at θ with index 0, i.e. $f(\lambda t)/f(t) \to_{t\to\theta} 1 \ \forall \lambda > 0$. Write $E(X \wedge t) =: L(t)$ and set $\varepsilon(t) := t(\log^+ L)'(t)$.

Both L and log are increasing and concave whence so is $\log L$, and $\varepsilon(t)/t$ decreases in t for t large.

By Karamata's representation theorem (Karamata, 1933, see also Bingham et al., 1987; Feller, 1966) $L(t) = E(X \wedge t)$ is slowly varying at ∞ iff $\varepsilon(t) \to 0$ as $t \to \infty$.

We will call an increasing function $A:\mathbb{R}_+\to\mathbb{R}_+$ *weakly regularly varying* if $\exists M>1$ such that $A(2t)\leqslant MA(t)$ and $2A(t)\leqslant A(Mt)$ $\forall$ large $t\in\mathbb{R}_+$. A decreasing function $B:\mathbb{R}_+\to\mathbb{R}_+$ will be so called if the increasing function $1/B$ is weakly regularly varying.

It can be shown that a monotone function $f:\mathbb{R}_+\to\mathbb{R}_+$ which is regularly varying at ∞ with non-zero index is weakly regularly varying, whereas a slowly varying function cannot be weakly regularly varying.

0.3. Dependence

The asymptotic behaviours (0.2)–(0.4) persist when the assumption of independence is relaxed to that of continued fraction mixing; the stationary process $(X_1,X_2,\ldots)$ being called *continued fraction mixing* (cf.-mixing) if $\vartheta(1)<\infty$ and $\vartheta(n)\downarrow 0$ where

$$\vartheta(n):=\sup\left\{\left|\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)\mathbb{P}(B)}-1\right|: A\in\sigma_1^k, B\in\sigma_{k+n}^\infty, \mathbb{P}(A)\mathbb{P}(B)>0, k\geqslant 1\right\}$$

where σ_k^N denotes the σ-algebra generated by the random variables $\{X_j: k\leqslant j<N+1\}$ $(k<N+1\leqslant\infty)$. The definition of $\vartheta(n)$ appears in Blum et al. (1963). See Bradley (1983) for the related concept of ψ-mixing.

Any probability preserving Gibbs–Markov map is cf.-mixing with $\vartheta(n)\downarrow 0$ exponentially (see Aaronson and Denker, 2001 or Section 4.7 of Aaronson, 1997).

The proof of (0.3) in the cf.-mixing case is the same as in the i.i.d. case, but uses the strong Borel–Cantelli lemma of Rényi (1970, p. 391). See Aaronson (1986) and Section 5 of Aaronson and Denker (1990) for (0.2); and Aaronson and Denker (1989) for (0.4).

0.4. Results

Let $(X_1,X_2,\ldots)$ be a non-negative, ergodic stationary process with $E(X\wedge t)=L(t)$. Set $a(t):=t/L(t)$ and $b:=a^{-1}$.

Write $\{X_k\}_{k=1}^n=\{r_{n,k}\}_{k=1}^n$, where $r_{n,1}\geqslant r_{n,2}\geqslant\cdots\geqslant r_{n,n}$ and set $M_n^{(v)}:=\sum_{k=1}^v r_{n,k}$. Let (for $r>0$) $J_r:=\sum_{n=1}^\infty \varepsilon(n)^r/n$ and define

$$\mathfrak{N}_X:=\begin{cases}\min\{\kappa\in\mathbb{N}: J_{\kappa+1}<\infty\} & \text{if } \exists\kappa,\ J_\kappa<\infty,\\ \infty, & \text{else.}\end{cases}$$

Note that $\mathfrak{N}_X<\infty$ implies that $L(t)$ is slowly varying at ∞ (as in this case $\varepsilon(t)\to 0$).

Theorem 1.1. (i) *Suppose that* $(X_1,X_2,\ldots)$ *is cf.-mixing, then*

$$\overline{\lim_{n\to\infty}}\sum_{k=1}^n 1_{[X_k>tb(n)]}=\mathfrak{N}_X\leqslant\infty\quad \forall t>0.$$

(ii) *Suppose that* $\sum_{n=1}^{\infty} \vartheta(n)/n < \infty$, *and that* $\mathfrak{N}_X < \infty$, *then* $\exists b_n = \mathrm{o}(b(n))$ (*depending only on the distribution of* X) *such that*

$$S_n - M_n^{(\mathfrak{N}_X)} \sim S_n^{(b_n)} \sim b(n) \quad a.s.\ as\ n \to \infty.$$

where $S_n^{(b)} := \sum_{k=1}^n X_k \wedge b$.

Remarks.

(1) It follows from (i) of Theorem 1.1, that $\overline{\lim}_{n\to\infty} (1/b(n))(S_n - M_n^{(K)}) = \infty$ a.s. $\forall K < \mathfrak{N}_X$ and it follows from (ii) of Theorem 1.1, that $(1/b(n))(S_n - M_n^{(K)}) \to 1$ a.s. $\forall K \geqslant \mathfrak{N}_X$.
(2) It is not hard to show using Birkhoff's theorem, that if $(X_1, X_2, \ldots)$ is an ergodic, stationary process with $E(|X|) < \infty$, then $(1/n)(S_n - M_n^{(K)}) \to E(X)$ a.s. $\forall K \in \mathbb{N}$.

In case $(X_1, X_2, \ldots)$ are i.i.d.r.v.'s, Theorem 1.1 follows from Theorem 1 in Mori (1977). The proof of Theorem 1.1 (given in Section 1) differs from that of Theorem 1 in Mori (1977) mainly in the estimation of large deviation probabilities of truncated sums. The use of log-moment generating functions in Mori (1977) is not possible here due to the dependence. We use moment estimations. Also the truncations are different.

In Section 2, we present examples of mixing, non-negative Markov chains $(X_1, X_2, \ldots)$ satisfying (0.2), (0.4), (0.3) and $\mathfrak{N}_X = 1$, but violating Theorem 1.1 in that

$$\overline{\lim_{n\to\infty}} \frac{1}{b(n)} (S_n - M_n^{(K)}) = \infty \quad \text{a.s. } \forall K \in \mathbb{N}.$$

Section 3 is an application of Theorem 1.1 to modified continued fractions. Let $x = \cfrac{1}{b_1 - \cfrac{1}{b_2 - \cfrac{1}{\ddots}}}$, then (as shown in Aaronson, 1986) $1/n \sum_{k=1}^n b_k \xrightarrow{\mathrm{P}} 3$ with respect to Lebesgue measure on $[0, 1]$. We show that $(1/n) \sum_{k=1}^n b_k \nrightarrow$ a.s.

1. Proof of Theorem 1.1

We will use the (elementary) fact that if $A : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing, weakly regularly varying, and $h(n) \downarrow$, $\gamma > 0$ then

$$\sum_{n=1}^{\infty} n^\gamma h(A(n)) < \infty \quad \text{implies} \quad \sum_{n=1}^{\infty} n^\gamma h(\varepsilon A(n)) < \infty \ \forall \varepsilon > 0$$

since if $K \in \mathbb{N}$ satisfies $\varepsilon A(Kn) \geqslant A(n)$, then

$$\sum_{n=1}^{\infty} n^\gamma h(\varepsilon A(n)) = \sum_{j=0}^{K-1} \sum_{n=1}^{\infty} (Kn + j)^\gamma h(\varepsilon A(Kn + j))$$

$$\leqslant K^{\gamma+1} \sum_{n=1}^{\infty} (n + 1)^\gamma h(A(n)) < \infty.$$

Let $N_{n,b} := \#\{k \leqslant n : X_k > b\}$ $(b > 0)$.

The following is a straightforward generalisation of Lemma 3 in Mori (1976) and Lemma 2 in Mori (1977) to the cf.-mixing case, and we only give a sketch of the proof.

Write $c(t) := P(X > t) = L'(t)$ where (as before) $E(X \wedge t) =: L(t)$.

Lemma 1.2. *Suppose that* $(X_1, X_2, \ldots)$ *is cf.-mixing and that* $B : \mathbb{R}_+ \to \mathbb{R}_+$ *is non-decreasing and satisfies* $nc(B(n)) \to 0$, *then for* $v \in \mathbb{N}$,

$$\overline{\lim_{n\to\infty}}\, N_{n,B(n)} \leqslant v \quad \text{a.s.} \quad \Leftrightarrow \quad \sum_{n=1}^{\infty} n^v P(X > B(n))^{v+1} < \infty.$$

In this case, if $B : \mathbb{R}_+ \to \mathbb{R}_+$ *is weakly regularly varying, then*

$$\overline{\lim_{n\to\infty}}\, N_{n,hB(n)} \leqslant v \quad \text{a.s.} \quad \forall h > 0.$$

Proof. As above,

$$\sum_{n=1}^{\infty} n^v P(X > B(n))^{v+1} \asymp \sum_{n=1}^{\infty} n^v P(X > hB(n))^{v+1} \quad \forall h > 0$$

in case $B : \mathbb{R}_+ \to \mathbb{R}_+$ is weakly regularly varying. The proof therefore splits into two parts:

$$P(N_{n,B(n)} \geqslant v) \asymp (nc(B(n)))^v \quad \forall v \geqslant 1 \tag{1.1}$$

and

$$\overline{\lim_{n\to\infty}}\, N_{n,B(n)} \leqslant v \quad \text{a.s.} \quad \Leftrightarrow \quad \sum_{n=1}^{\infty} \frac{P(N_{n,B(n)} \geqslant v+1)}{n} < \infty. \tag{1.2}$$

Set $N_n = N_{n,B(n)}$ for $n \geqslant 1$. To establish (1.1), suppose that M is as in the definition of cf.-mixing and that $\vartheta(\kappa) < 1$.

$$\begin{aligned} P(N_n \geqslant v) &\leqslant \sum_{K \subset \{1,\ldots,n\}, \#K = v} P(X_k > B(n)\ \forall k \in K) \\ &\leqslant M^v \binom{n}{v} c(B(n))^v \\ &\ll n^v c(B(n))^v. \end{aligned}$$

Now fix $n \geqslant \kappa$ so that $nc(B(n)) < \frac{1}{2}$. For $1 \leqslant k \leqslant n$ let

$$A_k := \bigcap_{1 \leqslant j \leqslant n, |j-k| \geqslant \kappa} [X_k > B(n), X_j \leqslant B(n)],$$

then

$$\sum_{k=1}^{n} 1_{A_k} \leqslant 2\kappa 1_{[N_n \geqslant 1]}$$

and

$$\begin{aligned} P(A_k) &\geqslant (1-\vartheta(\kappa))^2 P\left(\bigcap_{j=1}^{k-\kappa}[X_j \leqslant B(n)]\right) c(B(n)) P\left(\bigcap_{j=k+\kappa}^{n}[X_j \leqslant B(n)]\right) \\ &\geqslant (1-\vartheta(\kappa))^2(1-kc(B(n)))c(B(n))(1-(n-k)c(B(n))) \\ &\geqslant \frac{1}{4}(1-\vartheta(\kappa))^2 c(B(n)), \end{aligned}$$

whence

$$P(N_n \geqslant 1) \geqslant \frac{1}{2\kappa}\sum_{k=1}^{n} P(A_k) \geqslant \frac{1}{8\kappa}(1-\vartheta(\kappa))^2 nc(B(n)) =: \eta nc(B(n)).$$

It now follows that for $n \gg \nu\kappa$ so large that $nc(B(n)) < \frac{1}{2}$

$$\begin{aligned} P(N_n \geqslant \nu) &\geqslant P\left(\sum_{\ell=1}^{(n/\nu)-\kappa} 1_{[X_{j(n/\nu)+\ell} > B(n)]} \geqslant 1 \ \forall\ 0 \leqslant j \leqslant \nu-1\right) \\ &\geqslant (1-\vartheta(\kappa))^{\nu} P(N_{n/\nu-\kappa,B(n)} \geqslant 1)^{\nu} \geqslant (1-\vartheta(\kappa))^{\nu}\left(\eta\left(\frac{n}{\nu}-\kappa\right)c(B(n))\right)^{\nu} \\ &\gg n^{\nu} c(B(n))^{\nu}. \end{aligned}$$

This establishes (1.1).

The proof of (1.2) is that of Lemma 3 of Mori (1976), but using the strong Borel–Cantelli lemma of Rényi (1970, p. 391) which is valid for cf.-mixing processes instead of the classical one (which is only valid for i.i.d.r.v.'s). □

Proof of (i) of Theorem 1.1. By Lemma 1.2, a.s.,

$$\overline{\lim_{n\to\infty}} \sum_{k=1}^{n} 1_{[X_k > b(n)]} = \min\left\{\kappa \geqslant 1 : \sum_{n=1}^{\infty} n^{\kappa} c(b(n))^{\kappa+1} < \infty\right\}.$$

Using $c(x) = \varepsilon(x)L(x)/x$ and $b(n+1) - b(n) \asymp L(b(n)) = b(n)/n$, we have for $r > 0$,

$$\begin{aligned} \sum_{n=1}^{\infty} n^{r-1} c(b(n))^r &= \sum_{n=1}^{\infty} \frac{\varepsilon(b(n))^r}{n} \asymp \sum_{n=1}^{\infty} (b(n+1)-b(n)) \frac{\varepsilon(b(n))^r}{b(n)} \\ &\asymp \sum_{n=1}^{\infty} \sum_{b(n) \leqslant k < b(n+1)} \frac{\varepsilon(k)^r}{k} = J_r. \end{aligned}$$

Thus, $\min\{\kappa \geqslant 1 : \sum_{n=1} n^{\kappa} c(b(n))^{\kappa+1} < \infty\} = \mathfrak{N}_X$ establishing (i). □

Proof of (ii) of Theorem 1.1. The main ingredient here is the estimation of moments of truncated sums in Claim 1.

Define $\Delta(b) := \frac{1}{L(b)} \int_0^1 \varepsilon(bt)L(bt)\,\mathrm{d}t$, then $\Delta(b) \underset{b\to\infty}{\to} 0$.

As in Mori (1977) (but with Δ in place of ε), define

$$\phi(x) := \frac{a(x)}{\sqrt{\Delta(x)}}.$$

We claim that $\phi(x) \uparrow \infty$ as $x \uparrow \infty$. Indeed

$$\frac{1}{\phi(x)^2} = \frac{\Delta(x)}{a(x)^2} = \frac{L(x)}{x}\frac{1}{x^2}\int_0^x tc(t)\,\mathrm{d}t \downarrow 0.$$

Set $b_n := \phi^{-1}(n)$.

Claim 1.

$$E\left(\left|\frac{S_n^{(b_n)}}{b(n)} - 1\right|^Q\right) \ll \Delta(b_n)^{(Q+1)/2} + \frac{1}{n}\sum_{k=1}^n \vartheta(k) \quad \forall Q \in 2\mathbb{N}. \tag{1.4}$$

Proof. Fix $n \geqslant 1$ and set $Y_n := X \wedge b_n - L(b_n)$, then

$$E(|S_n^{(b_n)} - nL(b_n)|^Q) = E\left(\left(\sum_{k=1}^n Y_k\right)^Q\right) = \sum_{1\leqslant k_1,\ldots,k_Q\leqslant n} E\left(\prod_{i=1}^Q Y_{k_i}\right).$$

The latter sums need further organisation before estimation.

Given $1 \leqslant k_1,\ldots,k_Q \leqslant n$ let $K := \{k\in\mathbb{N} : \exists 1 \leqslant j \leqslant Q, k = k_j\} = \{\kappa_1,\ldots,\kappa_v\}$ where $v \leqslant Q$ and $\kappa_1 < \cdots < \kappa_v$, and define $f:\{1,\ldots,v\}\to\mathbb{N}$ by $f(j) := \#\{1\leqslant i\leqslant Q : k_i = \kappa_j\}$, then $\sum_{j=1}^v f(j) = Q$ and it follows that

$$\sum_{1\leqslant k_1,\ldots,k_Q\leqslant n} E\left(\prod_{i=1}^Q Y_{k_i}\right) = \sum_{v=1}^Q \sum_{1\leqslant \kappa_1<\cdots<\kappa_v\leqslant n} \sum_{f\in E_v^{(Q)}} E\left(\prod_{j=1}^v Y_{\kappa_j}^{f(j)}\right),$$

where

$$E_v^{(Q)} := \left\{ f:\{1,\ldots,v\}\to\mathbb{N}, \sum_{j=1}^v f(j) = Q = 2p \right\}.$$

There are two cases: $f \geqslant 2$ and $\min_{1\leqslant k\leqslant v} f(k) = 1$. Given $1 \leqslant v \leqslant Q$ let

$$F_v^{(Q)} := \{f\in E_v^{(Q)} : f \geqslant 2\}, \quad G_v^{(Q)} := \left\{ f\in E_v^{(Q)} : \min_{1\leqslant k\leqslant v} f(k) = 1 \right\}.$$

It follows that

$$E(|S_n^{(b_n)} - nL(b_n)|^Q) \leqslant \sum_{v=1}^{Q} \sum_{f\in E_v^{(Q)}} \sum_{1\leqslant \kappa_1<\cdots<\kappa_v\leqslant n} \left| E\left(\prod_{j=1}^{v} Y_{\kappa_j}^{f(j)}\right)\right|$$

$$= \sum_{v=1}^{Q} \sum_{f\in F_v^{(Q)}} + \sum_{v=1}^{Q} \sum_{f\in G_v^{(Q)}}. \tag{1.4}$$

Since $F_v^{(Q)} = \emptyset$ for $v > p$, we have by cf.-mixing that

$$\sum_{v=1}^{Q} \sum_{f\in F_v^{(Q)}} = \sum_{v=1}^{p} \sum_{f\in F_v^{(Q)}}$$

$$\ll \sum_{v=1}^{p} \sum_{f\in F_v^{(Q)}} \sum_{1\leqslant \kappa_1<\cdots<\kappa_v\leqslant n} \prod_{j=1}^{v} E(|Y_{\kappa_j}|^{f(j)}).$$

For $r \geqslant 2$ we have

$$E(|Y|^r) \leqslant 2^r E((X\wedge b_n)^r) = 2^r r \int_0^{b_n} x^{r-2}\varepsilon(x)L(x)\,\mathrm{d}x$$

$$= r2^r b_n^{r-1} \int_0^1 t^{r-2}\varepsilon(b_n t)L(b_n t)\,\mathrm{d}t = r2^r b_n^{r-1} L(b_n)\Delta(b_n),$$

so for $1 \leqslant \kappa_1 < \cdots < \kappa_v \leqslant n$ and $f \in F_v^{(Q)}$:

$$\prod_{j=1}^{v} E(|Y_{\kappa_j}|^{f(j)}) \ll \prod_{k=1}^{v} (b_n^{f(k)-1}\Delta(b_n)L(b_n)) = b_n^Q \left(\frac{L(b_n)\Delta(b_n)}{b_n}\right)^v.$$

Now

$$\frac{L(x)}{x} \sim \frac{1}{a(x)} = \frac{1}{\phi(x)\sqrt{\Delta(x)}}$$

whence

$$\frac{L(b_n)}{b_n} = \frac{1}{\phi(b_n)\sqrt{\Delta(b_n)}} = \frac{1}{n\sqrt{\Delta(b_n)}}$$

and

$$\prod_{j=1}^{v} E(|Y_{\kappa_j}|^{f(j)}) \ll b_n^Q \frac{\Delta(b_n)^{v/2}}{n^v}.$$

Thus,

$$\sum_{v=1}^{Q} \sum_{f\in F_v^{(Q)}} \ll \sum_{v=1}^{p} \binom{n}{v} b_n^Q \frac{\Delta(b_n)^{v/2}}{n^v} \asymp \sum_{v=1}^{p} b_n^Q \Delta(b_n)^{v/2} \sim b_n^Q \sqrt{\Delta(b_n)}.$$

We now turn to the estimation of $\sum_{f\in G_v^{(Q)}}$ in (1.4). Although $E(|X\wedge b_n|^r)=o(b_n^{r-1}L(b_n))\ \forall r\geqslant 2$, we have $E(|X\wedge b_n|)=L(b_n)$, which is too large, and we must use cf.-mixing more delicately in this case.

Fix $v\leqslant Q$, $f\in G_v^{(Q)}$ and suppose that $1\leqslant J\leqslant v$ satisfies $f(J)=1$. We will do the "generic" (difficult) case $2\leqslant J\leqslant v-1$ ($\Rightarrow v\geqslant 3$).

$$\sum_{1\leqslant\kappa_1<\cdots<\kappa_v\leqslant n}\left|E\left(\prod_{i=1}^{v}Y_{\kappa_i}^{f(i)}\right)\right|$$
$$=\sum_{L=1}^{n}\sum_{1\leqslant\kappa_1<\cdots<\kappa_{J-1}\leqslant L-1}\ \sum_{L+1\leqslant\kappa_{J+1}<\cdots<\kappa_v\leqslant n}\left|E\left(\prod_{i=1}^{J-1}Y_{\kappa_i}^{f(i)}Y_L\prod_{i=J+1}^{v}Y_{\kappa_i}^{f(i)}\right)\right|.$$

Fix $\kappa_1<\cdots<\kappa_{J-1}<L<\kappa_{J+1}<\cdots<\kappa_v\leqslant n$. By cf.-mixing and $E(Y_L)=0$,

$$\left|E\left(\prod_{i=1}^{J-1}Y_{\kappa_i}^{f(i)}Y_L\prod_{i=J+1}^{v}Y_{\kappa_i}^{f(i)}\right)\right|$$
$$\leqslant E\left(\prod_{i=1}^{J-1}|Y_{\kappa_i}|^{f(i)}\right)E(|Y_L|)E\left(\prod_{i=J+1}^{v}|Y_{\kappa_i}^{f(i)}|\right)$$
$$\times(\vartheta(L-\kappa_{J-1})+\vartheta(\kappa_{J+1}-L))$$
$$\ll b_n^{Q-v}L(b_n)^v(\vartheta(L-\kappa_{J-1})+\vartheta(\kappa_{J+1}-L)),$$

whence, by the above

$$\sum_{1\leqslant\kappa_1<\cdots<\kappa_v\leqslant n}\left|E\left(\prod_{i=1}^{v}Y_{\kappa_i}^{f(i)}\right)\right|$$
$$\ll b_n^{Q-v}L(b_n)^v\sum_{1\leqslant K<L<K'\leqslant n}\binom{K-1}{J-2}\binom{n-K'-1}{v-J-1}$$
$$\times(\vartheta(L-K)+\vartheta(K'-L))$$
$$\leqslant b_n^{Q-v}L(b_n)^v n^{v-3}\sum_{1\leqslant K<L<K'\leqslant n}(\vartheta(L-K)+\vartheta(K'-L))$$
$$\leqslant 2b_n^{Q-v}L(b_n)^v n^{v-3}n^2\sum_{k=1}^{n}\vartheta(k)$$
$$\ll n^{v-1}b_n^{Q-v}L(b_n)^v\sum_{k=1}^{n}\vartheta(k)=\frac{b_n^Q}{n}\left(\frac{1}{\Delta(b_n)}\right)^{v/2}\sum_{k=1}^{n}\vartheta(k).$$

It follows that

$$\sum_{v=1}^{Q}\sum_{f\in E_v^{(Q)}}\sum_{1\leqslant\kappa_1<\cdots<\kappa_v\leqslant n}\left|E\left(\prod_{k\in K}Y^{f(k)}\right)\right| \ll \frac{b_n^Q}{n}\sum_{v=1}^{Q}\left(\frac{1}{\Delta(b_n)}\right)^{v/2}\sum_{k=1}^{n}\vartheta(k)$$

$$\sim \frac{b_n^Q}{n}\left(\frac{1}{\Delta(b_n)}\right)^{Q/2}\sum_{k=1}^{n}\vartheta(k).$$

Putting things together

$$E(|S_n^{(b_n)}-nL(b_n)|^Q)\ll b_n^Q\left(\sqrt{\Delta(b_n)}+\frac{1}{n}\left(\frac{1}{\Delta(b_n)}\right)^{Q/2}\sum_{k=1}^{n}\vartheta(k)\right).$$

Next, note that $\phi(x)=a(x)/\sqrt{\Delta(x)}$ whence

$$a(\phi^{-1}(x))=x\sqrt{\Delta(\phi^{-1}(x))},\ \ a(b_n)=n\sqrt{\Delta(b_n)}$$

and

$$E\left(\left|\frac{S_n^{(b_n)}}{nL(b_n)}-1\right|^Q\right) \ll \left(\frac{b_n}{nL(b_n)}\right)^Q\left(\sqrt{\Delta(b_n)}+\frac{1}{n}\left(\frac{1}{\Delta(b_n)}\right)^{Q/2}\sum_{k=1}^{n}\vartheta(k)\right)$$

$$=\Delta(b_n)^{(Q+1)/2}+\frac{1}{n}\sum_{k=1}^{n}\vartheta(k)\to 0.$$

Thus $S_n^{(b_n)}/nL(b_n)\xrightarrow{\mathrm{P}}1$. Since $nc(b_n)\to 0$, we have $S_n/nL(b_n)\xrightarrow{\mathrm{P}}1$, whence $nL(b_n)\sim b(n)$ and

$$E\left(\left|\frac{S_n^{(b_n)}}{b(n)}-1\right|^Q\right)\ll\Delta(b_n)^{(Q+1)/2}+\frac{1}{n}\sum_{k=1}^{n}\vartheta(k)$$

which is (1.4) and the claim is established. □

Claim 2.

$$\sum_{n=1}^{\infty}\frac{1}{n}P\left(\left[\left|\frac{S_n^{(b_n)}}{b(n)}-1\right|>\varepsilon\right]\right)<\infty\quad\forall\varepsilon>0. \tag{1.5}$$

Proof. By the Chebyshev–Markov inequality,

$$P\left(\left[\left|\frac{S_n^{(b_n)}}{b(n)}-1\right|>\varepsilon\right]\right)\ll E\left(\left|\frac{S_n^{(b_n)}}{b(n)}-1\right|^Q\right),\quad \forall Q>1,$$

so by claim 1, (1.5) will follow from $\sum_{n=1}^{\infty}\Delta(b_n)^{(Q+1)/2}/n<\infty$ for some $Q>1$ and $\sum_{n=1}^{\infty}(1/n^2)\sum_{k=1}^{n}\vartheta(k)<\infty$. The latter follows from the assumptions on

$\{\vartheta(n)\}_{n\geqslant 1}$ as

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{k=1}^{n}\vartheta(k)=\sum_{k=1}^{\infty}\vartheta(k)\sum_{n=k}^{\infty}\frac{1}{n^2}\asymp\sum_{k=1}^{\infty}\frac{\vartheta(k)}{k}<\infty.$$

We will show that

$$\sum_{n=1}^{\infty}\frac{\Delta(b_n)^{\kappa}}{n}\asymp J_{\kappa}\quad\forall\kappa>0. \tag{1.6}$$

The proof of (1.6) is in two parts.

Firstly, for $\kappa,\ \gamma>0$ and writing $\gamma'=\phi(\gamma)$, we have

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\Delta(b_n)^{\kappa}}{n} &\asymp \int_{\gamma'}^{\infty}\frac{\Delta(\phi^{-1}(x))^{\kappa}\,\mathrm{d}x}{x}=\int_{\gamma'}^{\infty}\frac{a(\phi^{-1}(x)))^{2\kappa}\,\mathrm{d}x}{x^{2\kappa+1}}\\ &\xleftarrow{}_{t\to\infty}\int_{\gamma}^{t}\frac{a(y)^{2\kappa}\phi'(y)\,\mathrm{d}y}{\phi(y)^{2\kappa+1}}\\ &= \left[\frac{-a(y)^{2\kappa}}{2\kappa\phi(y)^{\kappa}}\right]_{\phi^{-1}(\gamma)}^{t}+\int_{\gamma}^{t}\frac{a(y)^{2\kappa-1}a'(y)\,\mathrm{d}y}{\phi(y)^{2\kappa}}\\ &= \int_{\gamma}^{t}\frac{L(y)\Delta(y)^{\kappa}a'(y)\,\mathrm{d}y}{y}+\mathrm{o}(1)\\ &\asymp \int_{\gamma}^{\infty}\frac{\Delta(y)^{\kappa}\,\mathrm{d}y}{y}.\end{aligned}$$

Next, we show that $\int_{\gamma}^{\infty}\Delta(y)^{\kappa}\,\mathrm{d}y/y\asymp J_{\kappa}$.

We start with $J_{\kappa}\ll\int_{c}^{\infty}\frac{\Delta(x)^{\kappa}\,\mathrm{d}x}{x}$ because $\varepsilon\ll\Delta$. To see this, recall that $\varepsilon(x)/x\downarrow$ whence $\varepsilon(by)\geqslant y\varepsilon(b)\ \forall b>0,\ 0<y<1$ and

$$\Delta(b)=\frac{1}{L(b)}\int_{0}^{1}\varepsilon(by)L(by)\,\mathrm{d}t\geqslant\frac{\varepsilon(b)}{L(b)}\int_{0}^{1}yL(by)\,\mathrm{d}t\sim\frac{\varepsilon(b)}{2}.$$

To show $\int_{c}^{\infty}\Delta(x)^{\kappa}\,\mathrm{d}x/x\ll J_{\kappa}$:

$$\begin{aligned}\int_{1}^{\infty}\frac{\Delta(b)^{\kappa}\,\mathrm{d}b}{b} &= \int_{1}^{\infty}\frac{1}{b}\left(\int_{0}^{1}\varepsilon(bt)\frac{L(bt)\,\mathrm{d}t}{L(b)}\right)^{\kappa}\mathrm{d}b\\ &\overset{\text{Jensen's ineq.}}{\leqslant}\int_{1}^{\infty}\frac{1}{b}\int_{0}^{1}\varepsilon(bt)^{\kappa}\frac{L(bt)\,\mathrm{d}t}{L(b)}\,\mathrm{d}b\end{aligned}$$

$$\begin{aligned}
&\leqslant \int_0^1 \int_1^\infty \frac{\varepsilon(bt)^\kappa \, \mathrm{d}b \, \mathrm{d}t}{b} \overset{y:=bt}{=} \int_0^1 \int_t^\infty \frac{\varepsilon(y)^\kappa \, \mathrm{d}y \, \mathrm{d}t}{y} \\
&= \int_1^\infty \frac{\varepsilon(y)^\kappa \, \mathrm{d}y}{y} + \int_0^1 \int_t^1 \frac{\varepsilon(y)^\kappa \, \mathrm{d}y \, \mathrm{d}t}{y} \\
&= \int_1^\infty \frac{\varepsilon(y)^\kappa \, \mathrm{d}y}{y} + \int_0^1 \varepsilon(y)^\kappa \, \mathrm{d}y \\
&= J_\kappa + \mathrm{O}(1),
\end{aligned}$$

(1.6) and claim 2 are established. □

Claim 3. $\frac{S_n^{(b_n)}}{b(n)} \to 1$ a.s.

Proof. From Claim 2 by condensation,

$$\sum_{j=1}^\infty P\left(\left[\left|\frac{S_{[\lambda^j]}^{(b_{[\lambda^j]})}}{b([\lambda^j])} - 1\right| > \varepsilon\right]\right) < \infty \quad \forall \varepsilon > 0, \ \lambda > 1$$

whence

$$\frac{S_{[\lambda^j]}^{(b_{[\lambda^j]})}}{b([\lambda^j])} \to 1 \quad \text{a.s. } \forall \lambda > 1.$$

By monotonicity, $\forall \lambda > 1$, a.s.,

$$\frac{1}{\lambda} = \lim_{j\to\infty} \frac{S_{[\lambda^{j-1}]}^{(b_{[\lambda^{j-1}]})}}{b([\lambda^j])} \leqslant \varliminf_{n\to\infty} \frac{S_n^{(b_n)}}{b(n)} \leqslant \varlimsup_{n\to\infty} \frac{S_n^{(b_n)}}{b(n)} \leqslant \lim_{j\to\infty} \frac{S_{[\lambda^{j+1}]}^{(b_{[\lambda^{j+1}]})}}{b([\lambda^j])} = \lambda \quad \text{a.s.}$$

showing that $\frac{S_n^{(b_n)}}{b(n)} \to 1$ a.s. □

Claim 4.

$$\varlimsup_{n\to\infty} \sum_{k=1}^n 1_{[X_k > b_n]} \leqslant 2\mathfrak{N} + 2 \quad \text{a.s.}$$

Proof. By Lemma 1.2, it suffices to show

$$\sum_{n=1}^\infty n^{2\mathfrak{N}+1} c(b_n)^{2\mathfrak{N}+2} < \infty.$$

For $\kappa = 2\mathfrak{N} + 2$,

$$\begin{aligned}
\sum_{n=1}^\infty n^{\kappa-1} c(b_n)^\kappa &= \sum_{n=1}^\infty \frac{1}{n}\left(\frac{n\varepsilon(b_n)}{a(b_n)}\right)^\kappa \ll \sum_{n=1}^\infty \frac{1}{n}\left(\frac{n\Delta(b_n)}{a(b_n)}\right)^\kappa \\
&= \sum_{n=1}^\infty \frac{\Delta(b_n)^{\kappa/2}}{n} \overset{(1.6)}{\asymp} J_{\kappa/2} = J_{\mathfrak{N}+1} < \infty. \qquad \square
\end{aligned}$$

Claim 5.

$$S_n - M_n^{(\mathfrak{N}_X)} \sim b(n) \quad \text{a.s.}$$

Proof. $\forall \eta > 0$, a.s. for n large

$$S_n - M_n^{(\mathfrak{N}_X)} = S_n^{(\eta b(n))} = S_n^{(b_n)} \pm (2\mathfrak{N} + 2)\eta b(n)$$

whence

$$1 - (2\mathfrak{N} + 2)\eta \leqslant \underline{\lim}_{n\to\infty} \frac{S_n - M_n^{(v)}}{b(n)} \leqslant \overline{\lim}_{n\to\infty} \frac{S_n - M_n^{(v)}}{b(n)} \leqslant 1 + (2\mathfrak{N} + 2)\eta.$$

This finishes the proof of Theorem 1.1. $\square$

Example. If $\varepsilon(t) \to 0$, $\varepsilon(t) = (1/(\log t)^{o(1)})$ as $t \to \infty$ (e.g.), then $\mathfrak{N}_X = \infty$.

If $\varepsilon(t) = o((1/\log\log\log t))$, then $L(t) \sim L(t \log\log t)$ and (0.4) holds.

Both conditions are satisfied for $L(t) = e^{\log(t+30)/\log\log(t+30)}$. Thus, there are processes (i.i.d.r.v.'s) $(X_1, X_2, \ldots)$ satisfying (0.4), but for which $\mathfrak{N}_X = \infty$ and trimming of any bounded number of maxima will not ensure a.s. convergence.

2. Markov chains with no trimmed Strong law

In this section, we construct examples showing that Theorem 1 fails for general mixing Markov chains.

Examples. There are non-negative, mixing Markov chains $(Y_1, Y_2, \ldots)$ satisfying $E(Y) = \infty$, $\mathfrak{N}_Y = 1$, (0.2)–(0.4) with normalising constants $b(n) = nE(Y \wedge b(n))$; but such that

$$\overline{\lim}_{n\to\infty} \frac{(S_n - M_n^{(K)})}{b(n)} = \infty \quad \text{a.s.} \quad \forall K \in \mathbb{N}.$$

For convenience, we construct the Markov chains over probability preserving transformations. Let S be an ergodic probability preserving transformation of the standard probability space $(\Omega, \mathscr{A}, P)$ and $f : \Omega \to \mathbb{N}$ be measurable, integrable and so that $\{f \circ S^n : n \geqslant 0\}$ are independent (e.g. $\Omega = \mathbb{N}^{\mathbb{N}}, S = \text{shift}, f(x) = x_1$ and P is a product measure).

Build $(X, \mathscr{B}, q, T)$ the tower transformation over S with height function f (see Kakutani, 1943 or Section 1.5 of Aaronson, 1997). This is an ergodic probability preserving transformation:

$$X := \{(x, n) : 1 \leqslant n \leqslant f(x)\}, \quad q(A \times \{n\}) := \frac{P(A)}{E(f)},$$

$$T(x, n) := \begin{cases} (x, n+1) & n < f(x)), \\ (Sx, 1) & n = f(x). \end{cases}$$

Now define $g : X \to \mathbb{N}$ by $g(x, n) := n$.

Our examples will be of form $(Y_1, Y_2, \ldots) := (g, g \circ T, g \circ T^2, \ldots)$. A calculation indeed shows that the ergodic stationary process $(g, g \circ T, g \circ T^2, \ldots)$ is a Markov chain (a renewal process) whose joint distributions are given by

$$q([g = s_0, g \circ T = s_1, \ldots, g \circ T^n = s_n]) = \pi_{s_0} p_{s_0,s_1} \ldots p_{s_{n-1},s_n}$$

where $\pi_s := P([f \geqslant s])/E(f)$ and

$$p_{j,k} = \begin{cases} \frac{P([f=j])}{E(f)\pi_j} & \text{if } j \in \mathbb{N},\ k = 1, \\ \frac{\pi_{j+1}}{\pi_j} & \text{if } j \in \mathbb{N},\ k = j+1, \\ 0, & \text{else.} \end{cases}$$

This chain is mixing if (e.g.) $P([f = n]) > 0\ \forall n \geqslant 1$ large.

Proposition 2.1 (Tanny, 1974).

$$\frac{g \circ T^n}{n} \underset{n\to\infty}{\to} 0 \quad \text{a.s.}$$

Proof. Since $E(f) < \infty$, we have $f \circ S^n/n \to 0$ a.s. on Ω. Next, for a.e. $x \in \Omega$ and $\forall n$ large, $\exists 0 \leqslant k_n \leqslant n$ such that $g(T^n x) \leqslant f(S^{k_n} x)$ whence $g \circ T^n/n \to 0$ a.s. on Ω. The proposition follows from the T-invariance of $\overline{\lim}_{n\to\infty}\, g \circ T^n/n$. $\square$

Next, we investigate the asymptotic behaviour of $g_n = g_n^{(T)} := \sum_{k=0}^{n-1} g \circ T^k$. To this end, let

$$\mathscr{L}(t) := E\left(\left(\frac{f(f+1)}{2}\right) \wedge t\right).$$

Lemma 2.2. (i) *If $\mathscr{L}(t)$ is slowly varying at ∞ and $E(f^2) = \infty$, then*

$$\mathscr{L}(t) \sim \tfrac{1}{2} E(f^2 \wedge t) \quad \text{as } t \to \infty.$$

(ii) *If $P([f \geqslant u]) \sim h(u)/u^2$ where $\int_1^\infty (h(u)\,\mathrm{d}u/u) = \infty$ and h is slowly varying at ∞, then $E(g) = \infty$, $\mathscr{L}$ is slowly varying at ∞ and*

$$L_g(t) := E(g \wedge t) \sim \frac{1}{E(f)} \mathscr{L}(t^2) \quad \text{as } t \to \infty.$$

Proof.

$$\text{(i)}\quad \frac{1}{2} E(f^2 \wedge t) \leqslant E\left(\frac{f^2}{2} \wedge t\right) \leqslant \mathscr{L}(t) \sim \mathscr{L}\left(\frac{t}{2}\right) = \frac{1}{2} E(f(f+1) \wedge t)$$

$$\sim \frac{1}{2} E(f^2 \wedge t).$$

To establish (ii), we first note that $\forall\varepsilon > 0$,

$$\int_1^t \frac{h(u)\,\mathrm{d}u}{u} \geqslant \int_{\varepsilon t}^t \frac{h(u)\,\mathrm{d}u}{u} \sim h(t)\log\frac{1}{\varepsilon} \quad \text{as } t\to\infty,$$

whence

$$h(t) = \mathrm{o}\left(\int_1^t \frac{h(u)\,\mathrm{d}u}{u}\right) \quad \text{as } t\to\infty.$$

It follows that $\int_1^t (h(u)\,\mathrm{d}u)/u$ is slowly varying at ∞ (because $\int_t^{\lambda t}(h(u)\,\mathrm{d}u/u) \sim h(t)\log\lambda$ as $t\to\infty$). Next

$$\frac{1}{2}E(f^2\wedge t) = \frac{1}{2}E((f\wedge\sqrt{t})^2) = \int_0^{\sqrt{t}} sP([f\geqslant s])\,\mathrm{d}s \sim \int_1^{\sqrt{t}} \frac{h(u)\,\mathrm{d}u}{u}$$

which latter is slowly varying at ∞. Analogously to the proof of (1), we see that $\mathscr{L}(t)$ is slowly varying at ∞. Next,

$$q(g\geqslant u) = \frac{1}{E(f)}\sum_{v=u}^{\infty} P(f\geqslant v) \sim \frac{h(u)}{E(f)u},$$

whence

$$L_g(t) = \sum_{k=1}^{t} q(g\geqslant k) \sim \frac{1}{E(f)}\sum_{u=1}^{t}\frac{h(u)}{u} \sim \frac{1}{E(f)}\mathscr{L}(t^2). \qquad \square$$

We use the notation $g_n = g_n^{(T)} := \sum_{k=0}^{n-1} g\circ T^k$.

Proposition 2.3.

(i) *Suppose that* $E(g)=\infty$, $\mathscr{L}$ *is slowly varying and let* $\beta(n) = n\mathscr{L}(\beta(n))$, *then*

$$\frac{g_n}{\beta(n)} \xrightarrow{q} \frac{1}{E(f)}, \qquad \overline{\lim_{n\to\infty}}\, \frac{g_n}{\beta(n)} = \infty \quad \text{a.s.}$$

and, in case $\mathscr{L}(t)\sim\mathscr{L}(t\log\log t)$:

$$\underline{\lim}_{n\to\infty} \frac{g_n}{\beta(n)} = \frac{1}{E(f)} \text{ a.s.}$$

(ii) *Under the assumptions of Lemma* 2.2 *and* $\mathscr{L}(t^2)\sim\mathscr{L}(t)$; $(g, g\circ T, \ldots)$ *satisfies* (0.2)–(0.4).

Proof. Note that $T_\Omega = T^f = S$ whence $T_\Omega^n = T^{f_n^{(S)}}$ where $f_n = f_n^{(S)} := \sum_{k=0}^{n-1} f\circ S^k$. It follows that on Ω:

$$g_{f_n^{(S)}}^{(T)} = h_n^{(S)},$$

where

$$h := g_f^{(T)} = \sum_{k=0}^{f-1} g \circ T^k = \frac{f(f+1)}{2}.$$

Since $\{h \circ S^n : n \geqslant 1\}$ are independent, by (0.2)–(0.4):

$$\frac{h_n^{(S)}}{\beta(n)} \xrightarrow{q} 1, \qquad \overline{\lim_{n\to\infty}} \frac{h_n^{(S)}}{\beta(n)} = \infty \quad \text{a.s.}$$

and, in case $\mathscr{L}(t) \sim \mathscr{L}(t \log\log t)$:

$$\underline{\lim_{n\to\infty}} \frac{h_n^{(S)}}{\beta(n)} = 1 \quad \text{a.s.}$$

By Birkhoff's pointwise ergodic theorem, $f_n \sim E(f)n$ a.s. on Ω, whence, a.s. on Ω (!):

$$\frac{g_{E(f)n}}{\beta(n)} \xrightarrow{q} 1, \ \overline{\lim_{n\to\infty}} \frac{g_{E(f)n}}{\beta(n)} = \infty \quad \text{and, in case } \mathscr{L}(t) \sim \mathscr{L}(t \log\log t),$$

$$\underline{\lim_{n\to\infty}} \frac{g_{E(f)n}}{\beta(n)} = 1.$$

Using the 1-regular variation of $\beta(n)$, and ergodicity of T, we establish (i) from which (ii) follows since $\mathscr{L}(t^2) \sim \mathscr{L}(t)$ implies $\beta(n) \sim E(f)b(n)$ where $b(n) = nE(g \wedge b(n))$. $\square$

Remark. Note that $\mathscr{L}(t^2) \sim \mathscr{L}(t)$ if $\varepsilon(t) := t(\log^+ L)'(t) = o(\frac{1}{\log t})$ as $t \to \infty$.

Proposition 2.4. *If* $E(g) = \infty$, *then*

$$\overline{\lim_{n\to\infty}} \frac{(g_n^{(T)} - M_n^{(K)})}{\beta(n)} = \infty \quad \text{a.s.} \quad \forall K \in \mathbb{N}.$$

Proof. $r_{n,1}(x) = g \circ T^{k_n(x)}(x)$ for some $0 \leqslant k_n(x) \leqslant n-1$. Thus,

$$M_n^{(K)} \leqslant K r_{n,1}(x) = Kg \circ T^{k_n(x)}(x) = o(n)$$

as $n \to \infty$ by Proposition 2.1. On the other hand, $E(g) = \infty$, so $g_n/n \to \infty$ and $M_n^{(K)} = o(g_n)$ a.s. $\square$

The advertised examples. If $P([f \geqslant t]) \sim h(t)/t^2$ as $t \to \infty$ where $1/h(t) = \prod_{j=1}^r \log^j(t+e_j)$ for some $r \in \mathbb{N}$ where $e_1 := e$, $e_{j+1} := e^{e_j}$, then $\mathscr{L}(t) \sim \log^{r+1}(t)$ $\sim \mathscr{L}(t^2)$ as $t \to \infty$ where $\log^1(t) := \log(t)$ and $\log^{r+1}(t) := \log(\log^r(t))$.

Thus, $E(g) = \infty$, $\mathfrak{N}_g = 1$, and $(g, g \circ T, \ldots)$ satisfies (0.2)–(0.4) with normalising constants $b(n) = nE(Y \wedge b(n))$ but $\overline{\lim}_{n\to\infty} (g_n^{(T)} - M_n^{(K)})/b(n) = \infty$ a.s. $\forall K \in \mathbb{N}$.

3. Applications

3.1. Modified continued fractions

Let

$$x=\cfrac{1}{b_1-\cfrac{1}{b_2-\cfrac{1}{\ddots}}},$$

then $b_n(x)=[1/(V^{n-1}x)]+1$ where $Vx:=1-\{1/x\}$. The transformation $V:[0,1]\to[0,1]$ has an infinite, invariant measure μ with density $(\mathrm{d}\mu/\mathrm{d}m)(x)=1/(1-x)$ with respect to which the function $b(x)=[1/x]+1$ is not integrable. Nevertheless (as shown in Aaronson, 1986)

$$A(n):=\frac{1}{n}\sum_{k=1}^{n} b_k \xrightarrow{\mathrm{P}} 3.$$

We prove here that a.s.,

$$\varliminf_{n\to\infty} A(n)=2, \quad \text{and} \quad \varlimsup_{n\to\infty} A(n)=\infty. \tag{3.1}$$

As shown in Dajani and Kraaikamp (2000)

$$A\left(\sum_{k=1}^{n} a_{2k-1}\right)=2+\frac{\sum_{k=1}^{n} a_{2k}}{\sum_{k=1}^{n} a_{2k-1}}$$

where $x=1/a_1+1/a_2+1/\ldots$. The regular continued fraction process $(a_1,a_2,\ldots)$ is given by $a_n(x):=a(U^{n-1}x)$ where $a(x):=[1/x]$ and $U:(0,1)\to(0,1)$ is defined by $Ux:=\{1/x\}$. Gauss' measure $\mathrm{d}\mathbb{P}(x)=\mathrm{d}x/(\log 2(1+x))$ is U-invariant on $[0,1]$. As shown in Doeblin (1940), it is cf.-mixing with $\vartheta(n)=\mathrm{O}(\theta^n)$ for some $0<\theta<1$.

Theorem 1.1 holds with $\mathfrak{N}_a=1$. The trimmed strong law for the regular continued fraction process was first established in Diamond and Vaaler (1986).

Thus, (3.1) follows from the following lemma.

Lemma 3.1. *Let $\{X_k\}_{k\geqslant 1}$ be a non-negative, stationary process with $\sum_{k=1}^{\infty}\vartheta(k)/k<\infty$, and suppose that $\mathfrak{N}_X<\infty$, then for $d\geqslant 2$ and $0\leqslant i\neq j<d$,*

$$\varliminf_{n\to\infty}\frac{\sum_{k=1}^{n} X_{dk+i}}{\sum_{k=1}^{n} X_{dk+j}}=0 \quad \text{and} \quad \varlimsup_{n\to\infty}\frac{\sum_{k=1}^{n} X_{dk+i}}{\sum_{k=1}^{n} X_{dk+j}}=\infty \quad \text{a.s.}$$

Proof. Since $\mathfrak{N}_X<\infty$, L is slowly varying at ∞, whence $b(t)$ defined by $b(t)=tL(b(t))$ is regularly varying at ∞ with index 1 . We claim first that $\exists\beta_n=\mathrm{o}(b(n))$ such that $\varlimsup_{n\to\infty}\sum_{k=1}^{n} 1_{[Z_k>\beta_n]}=\mathfrak{N}_X$ a.s. for any stationary process $\{Z_n\}$ with $\sum_{n=1}^{\infty}\vartheta(n)/n<\infty$ and $\operatorname{dist} Z=\operatorname{dist} X$.

By Lemma 1.2, $\sum_{n\in\mathbb{N}} n^{\mathfrak{N}_X} c(b(n)/k)^{\mathfrak{N}_X+1} < \infty$. To obtain such a sequence $\{\beta_n\}$, fix $m_k \uparrow$ such that

$$\sum_{n\geqslant m_k} n^{\mathfrak{N}_X} c\left(\frac{b(n)}{k}\right)^{\mathfrak{N}_X+1} < \frac{1}{2^k} \quad \forall k \geqslant 1$$

and set $\beta_n := b(n)/k$ for $n\in\mathbb{N}$, $m_k \leqslant n < m_{k+1}$. Evidently, $\beta_n = o(b(n))$ and $\sum_{n\in\mathbb{N}} n^{\mathfrak{N}_X} c(\beta_n)^{\mathfrak{N}_X+1} < \infty$, whence $\overline{\lim}_{n\to\infty} \sum_{k=1}^n 1_{[Z_k>\beta_n]} = \mathfrak{N}_X$ a.s.

By Theorem 1.1, $S_n^{(\beta_n)} \sim b(n)$ a.s., and to see

$$\overline{\lim_{n\to\infty}} \frac{\sum_{k=1}^n X_{dk+i}}{\sum_{k=1}^n X_{dk+j}} = \infty \quad \text{a.s.,}$$

fix $M > 0$ large and note that a.s., $\exists n_\ell \to \infty$ and $B_\ell \subset \{dk+i\}_{k=1}^{n_\ell}$, $|B_\ell| = \mathfrak{N}_X$ such that

(i) $X_k > Mb(n_\ell)$ $\forall k \in B_\ell$, and (ii) $X_k \leqslant \beta_{n_\ell}$ $\forall k \notin B_\ell$, $k \leqslant (d+1)n_\ell$. It follows that

$$\sum_{k=1}^{n_\ell} X_{dk+j} = \sum_{k=1}^{n_\ell} X_{dk+j} \wedge \beta_{n_\ell} \sim b(n_\ell) \quad \text{a.s.,}$$

whereas

$$\sum_{k=1}^{n_\ell} X_{dk+i} \geqslant M\mathfrak{N}_X b(n_\ell)$$

with the conclusion that

$$\overline{\lim_{n\to\infty}} \frac{\sum_{k=1}^n X_{dk+i}}{\sum_{k=1}^n X_{dk+j}} \geqslant \overline{\lim_{\ell\to\infty}} \frac{\sum_{k=1}^{n_\ell} X_{dk+i}}{\sum_{k=1}^{n_\ell} X_{dk+j}} \geqslant \lim_{\ell\to\infty} \frac{M\mathfrak{N}_X b(n_\ell)}{\sum_{k=1}^{n_\ell} X_{dk+j} \wedge \beta_{n_\ell}} = M\mathfrak{N}_X. \qquad \square$$

3.2. *Visits to cusps*

Define $W:[0,1]\to[0,1]$ by $W(x)=x/(1-x)$ $(0<x<\frac{1}{2})$ and $W(1-x)=1-W(x)$.

The measure $v \sim m$ with $\mathrm{d}v/\mathrm{d}m(x) = 1/(x(1-x))$ is W-invariant, and as shown in Thaler (1983) (see also Aaronson, 1997), $([0,1], m, W)$ is conservative and ergodic.

The invariant measure density v has "cusps" at 0 and 1 in the sense $\mu([0,\varepsilon)) = \mu([1-\varepsilon,1)) = \infty$ $\forall \varepsilon > 0$, but $\mu((a,b)) < \infty$ $\forall 0 < a < b < 1$ and it is natural to ask about the frequency of visits to these "cusps".

It was shown in Thaler (2000) that

$$\frac{1}{n}\sum_{k=0}^{n-1} 1_{[0,1/2)} \circ W^k \xrightarrow{m} \frac{1}{2}, \quad \text{whence } \frac{\sum_{k=0}^{n-1} 1_{[0,1/2)} \circ W^k}{\sum_{k=0}^{n-1} 1_{[1/2,1)} \circ W^k} \xrightarrow{m} 1. \tag{3.2}$$

We show, using (3.1) that

$$\underline{\lim_{n\to\infty}} \frac{\sum_{k=0}^{n-1} 1_{[0,1/2)}(W^k x)}{\sum_{k=0}^{n-1} 1_{[1/2,1)}(W^k x)} = 0, \qquad \overline{\lim_{n\to\infty}} \frac{\sum_{k=0}^{n-1} 1_{[0,1/2)}(W^k x)}{\sum_{k=0}^{n-1} 1_{[1/2,1)}(W^k x)} = \infty \tag{3.3}$$

(cf. Inoue, 1997, 2001).

Define $K:[0,1]\to\mathbb{Z}_+$ by $K(x):=\min\{j\geqslant 0: W^jx>\frac{1}{2}\}$ and $\tilde{W}:[0,1]\to[0,\frac{1}{2}]\times\{0,1\}$ by $\tilde{W}(x):=W^{K(x)+1}(x)$. It turns out that $K(x)=b(x)-2:=[1/x]-1$, $W(x)=V(x):=1-\{1/x\}$ (b,V as above), whence by (3.1), $\underline{\lim}_{n\to\infty}(K_n(x)/n)=0$ and $\overline{\lim}_{n\to\infty}(K_n(x)/n)=\infty$ a.s. where $K_n:=\sum_{k=0}^{n-1}K\circ V^k$.

This proves (3.3) as

$$\sum_{k=0}^{K_n(x)-1}1_{[0,1/2)}(W^kx)=K_n(x)\quad\text{and}\quad\sum_{k=0}^{K_n(x)-1}1_{[1/2,1)}(W^kx)=n.$$

Acknowledgements

The authors would like to thank the Institute of Mathematics of Nicholas Copernicus University, Toruń, Poland for hospitality provided when this project was initiated.

References

Aaronson, J., 1986. Random f-expansions. Ann. Probab. 14, 1037–1057.

Aaronson, J., 1977. On the ergodic theory of non-integrable functions and infinite measure spaces. Israel J. Math. 27, 163–173.

Aaronson, J., 1997. An Introduction to Infinite Ergodic Theory. In: Mathematical Surveys and Monographs, Vol. 50. American Mathematical Society, Providence, RI, US.

Aaronson, J., Denker, M., 1990. Upper bounds for ergodic sums of infinite measure preserving transformations. Trans. Amer. Math. Soc. 319, 101–138.

Aaronson, J., Denker, M., 1989. Lower Bounds for Partial Sums of Certain Positive Stationary Processes. Almost Everywhere Convergence (Columbus, OH, 1988). Academic Press, Boston, MA, pp. 1–9.

Aaronson, J., Denker, M., 2001. Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps. Stochastics Dynamics 1 (2), 193–237.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. In: Encyclopedia of Mathematics and its Applications, Vol. 27. Cambridge University Press, Cambridge, UK.

Blum, J.R., Hanson, D.L., Koopmans, L.H., 1963. On the strong law of large numbers for a class of stochastic processes. Z. Wahrsch. Verw. Gebiete 2, 1–11.

Bradley, R.C., 1983. On the ψ-mixing condition for stationary random sequences. Trans. Amer. Math. Soc. 276 (1), 55–66.

Chow, Y.S., Robbins, H., 1961. On sums of independent random variables with ∞ moments. Proc. Nat. Acad. Sci. USA 47, 330–335.

Dajani, K., Kraaikamp, C., 2000. The mother of all continued fractions (Dedicated to the memory of Anzelm Iwanik). Colloq. Math. 84/85 part 1, 109–123.

Diamond, H., Vaaler, J.D., 1986. Estimates for partial sums of continued fraction partial quotients. Pacific J. Math. 122 (1), 73–82.

Doeblin, W., 1940. Remarques sur la thorie mtrique des fractions continues. Compositio Math. 7, 353–371.

Feller, W., 1966. An Introduction to Probability Theory and its Applications, Vol. II. Wiley, New York, 1966.

Feller, W., 1945. Note on the law of large numbers and "fair" games. Ann. Math. Statist. 16, 301–304.

Feller, W., 1946. A limit theorem for random variables with infinite moments. Amer. J. Math. 68, 257–262.

Inoue, T., 1997. Ratio ergodic theorems for maps with indifferent fixed points. Ergodic Theory Dynamical Systems 17 (3), 625–642.

Inoue, T., 2001. Correction. Ergodic Theory Dynamical Systems 21 (4), 1273.

Kakutani, S., 1943. Induced measure preserving transformations. Proc. Imp. Acad. Sci. Tokyo 19, 635–641.

Karamata, J., 1933. Sur un mode de croissance reguliere. Theoreme s fondamentaux. Bull. Soc. Math. France 61, 55–62.

Klass, M., Teicher, H., 1977. Iterated logarithm laws for asymmetric random variables barely with or without finite mean. J. Ann. Probab. 5, 861–874.

Mori, T., 1976. The strong law of large numbers when extreme terms are excluded from sums. Z. Wahrsch. Verw. Gebiete 36 (3), 189–194.

Mori, T., 1977. Stability for sums of i.i.d. random variables when extreme terms are excluded. Z. Wahrsch. Verw. Gebiete 40 (2), 159–167.

Rényi, A., 1970. Probability Theory. North-Holland, Amsterdam.

Tanny, D., 1974. A 0–1 law for stationary sequences. Z. Wahrsch. Verw. Gebiete 30, 139–148.

Thaler, M., 1983. Transformations on [0, 1] with infinite invariant measures. Israel J. Math. 46, 67–96.

Thaler, M., 2000. A limit theorem for sojourn times near indifferent fixed points of one-dimensional maps, preprint.